Notes on sequential and repeated games

1 Sequential Move Games

Thus far we have examined games in which players make moves simultaneously (or without observing what the other player has done). Using the normal (strategic) form representation of a game we can identify sets of strategies that are best responses to each other (Nash Equilibria). We now focus on sequential games of complete information. We can still use the normal form representation to identify NE but sequential games provide more information than what is present in the normal form because some players observe other players' decisions before they take action. The fact that some actions are observable may cause some NE of the normal form representation to be inconsistent with what one might think a player would do.

Here's a simple game between an Entrant and an Incumbent. The Entrant moves first and the Incumbent observes the Entrant's action and then gets to make a choice. The Entrant has to decide whether or not he will enter a market or not. Thus, the Entrant's two strategies are "Enter" or "Stay Out". If the Entrant chooses "Stay Out" then the game ends. The payoffs for the Entrant and Incumbent will be 0 and 2 respectively. If the Entrant chooses "Enter" then the Incumbent gets to choose whether or not he will "Fight" or "Accommodate" entry. If the Incumbent chooses "Fight" then the Entrant receives -3 and the Incumbent receives -1. If the Incumbent chooses "Accommodate" then the Entrant receives 2 and the Incumbent receives 1. This game in normal form is

		Incumbent	
		Fight if Enter	Accommodate if Enter
Entrant	Enter	$-3, -1$	$2, 1$
	Stay Out	$0, 2$	$0, 2$

Note that there are two pure strategy Nash Equilibria (PSNE) to this game. One is that the Entrant chooses Enter and the Incumbent chooses Accommodate and the other is that the Entrant chooses Stay Out and the Incumbent chooses Fight.[1] Of these two PSNE, which seems more "believable"? The NE where the Entrant chooses Stay Out and the Incumbent chooses Fight is only a NE if the Entrant thinks that the Incumbent's choice of Fight is credible. But what does the Entrant know? The Entrant knows that if he chooses Enter that the Incumbent will not choose Fight but will choose Accommodate. If you are the Entrant, are you worried about the Incumbent choosing Fight (as this game is structured)? No, because if you choose Enter, the best thing for the Incumbent to do for himself at that point would be to choose Accommodate. Thus, we can rule out the NE of Stay Out, Fight because the choice of Fight is not credible (Note: this does NOT mean that Stay Out, Fight is NOT a NE, it just means that it relies on a noncredible threat).

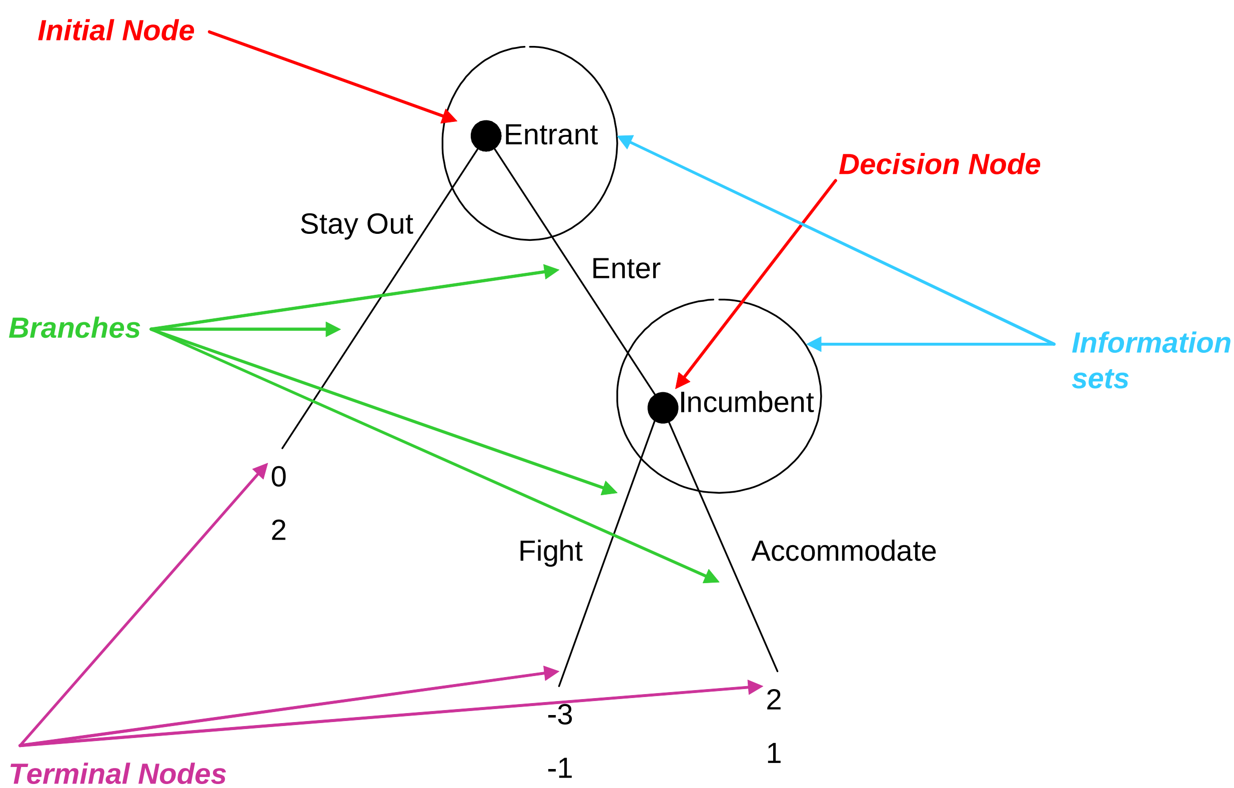
1.1 Extensive Form Representation

We can represent the sequential game using the extensive form representation or game tree. The extensive form representation, Γ_E, has more components than the normal form representation. Recall that the normal form representation required that we only need to know how many players there were, which strategies were available to each player, and which payoffs occurred as a result of the players' strategy choices. With an extensive form game we also need to consider the fact that players move at different points in time. An extensive form game will consist of the following basic items: players, decision nodes, information sets, strategies, and payoffs.[2] Note that the components of a normal form game are all here, so that any extensive

[1] To be complete, there is also a mixed strategy Nash Equilibrium (MSNE) where the Entrant chooses Stay Out with probability 1 and the Incumbent chooses Fight with probability $\frac{2}{5}$ and Accommodate with probability $\frac{3}{5}$.

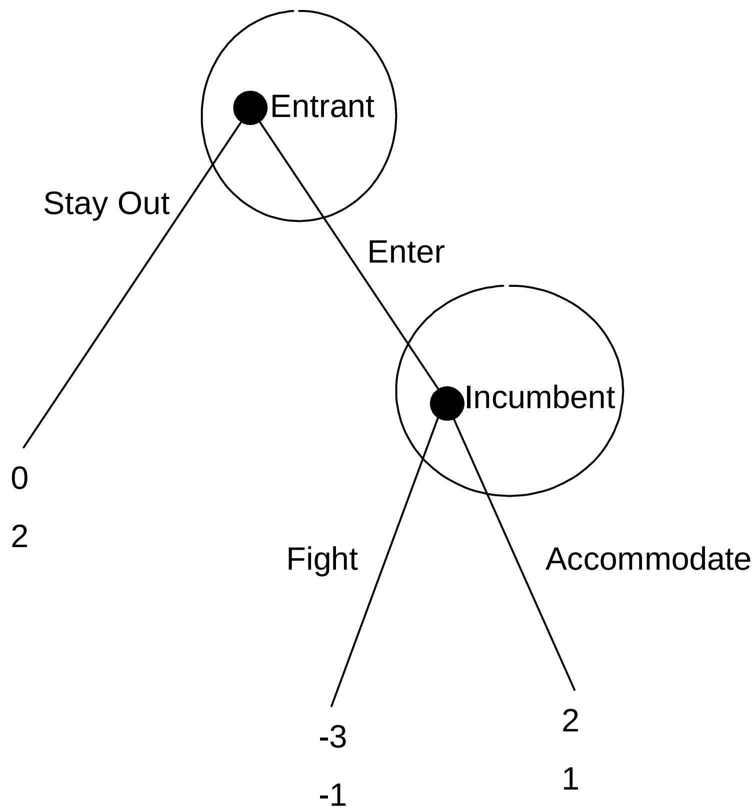
[2] There is a much more formal definition on page 227 of MWG.

form game may also be represented as a normal form game. The additional items are decision nodes and information sets. A decision node is a point where a player makes a decision. An information set is what the player knows at a certain node. The game begins with a single decision node, called the initial node. The player who makes a decision at a node is listed and the strategies available to that player are then drawn as lines from that node (the branches of the game tree). If the game ends after a player makes a decision, then the players reach a terminal node and payoffs are listed. The payoffs correspond to the path played out by the strategy choices that lead to the terminal node. If the game continues after a player makes a decision then the players reach another decision node. A different player will be able to take an action, and his strategies will be represented as branches extending from that node. This process continues until the game reaches a terminal node after strategy choices are made. The game tree representation of the Entrant, Incumbent game is in Figure 1.1. Note that an actual game tree does not include all the labels, but I have included them for reference.



Game tree with its components labeled.

The actual game tree, without the labels, would look like Figure 1.1.



Game tree without the components labeled.

Much simpler, but note that the players, strategies, and payoffs are still listed. You might ask why we circle the node and label it an information set. It is possible that players do not know which node they are at, so that the player's information set contains multiple nodes. Thus, the information set would be a circle around both nodes. Consider the simultaneous move Prisoner's Dilemma game. Prisoner 2 does not know which choice is made by Prisoner 1, so his information set contains both nodes, as in Figure 1.Contrast this with the sequential version of the Prisoner's Dilemma game where Prisoner 2 knows what Prisoner 1 has chosen in Figure 2.The games look similar, but they are slightly different as we will discuss shortly.

There is one other small detail in extensive form games. It may be that one or more players has an infinite number of strategies. Consider a game in which players may choose to produce any quantity of an item greater than or equal to 0 or less than or equal to the quantity consumers would purchase when price equals 0. The strategy space is then any real number in the interval $[0, Q_0]$, where Q_0 is the quantity consumers would purchase when price equals 0. As we cannot represent the strategies with a finite number of branches, we would use a dashed line between two branches to represent an interval. The branches would be labeled 0 and Q_0. If the other player does not know the choice of the first player (the game is simultaneous) then both nodes extending from the two branches as well as all the nodes represented by the dashed line are in the information set, so we circle both nodes and the dashed line. If the second player does observe the first player's choice, then we simply draw a circle around some elements of the dashed line but not the two nodes drawn from the branches.

Representing an extensive form game in normal form We know that with a normal form game we only need to know the players, strategies, and payoffs. Representing Entry, Incumbent in normal form is easy: because there are only 2 players and 2 strategies we only need a 2x2 matrix. We have already seen the representation of the simultaneous Prisoner's Dilemma. But what does the normal form representation of the sequential Prisoner's Dilemma look like? There are 2 players, Prisoner 1 and Prisoner 2. Prisoner 1 has 2 strategies Confess and Not Confess. So far everything looks the same. But Prisoner 2 now has four strategies. How can that be? Prisoner 2 only chooses Confess or Not Confess, doesn't he? While that is

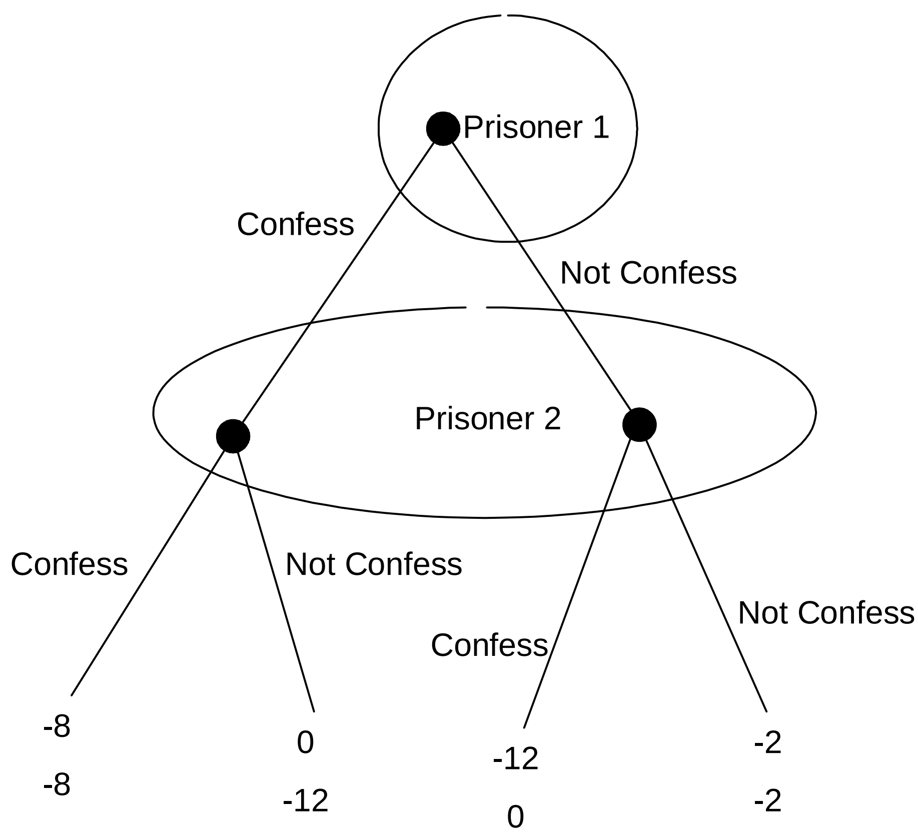


Figure 1: Simultaneous move Prisoner's Dilemma game.

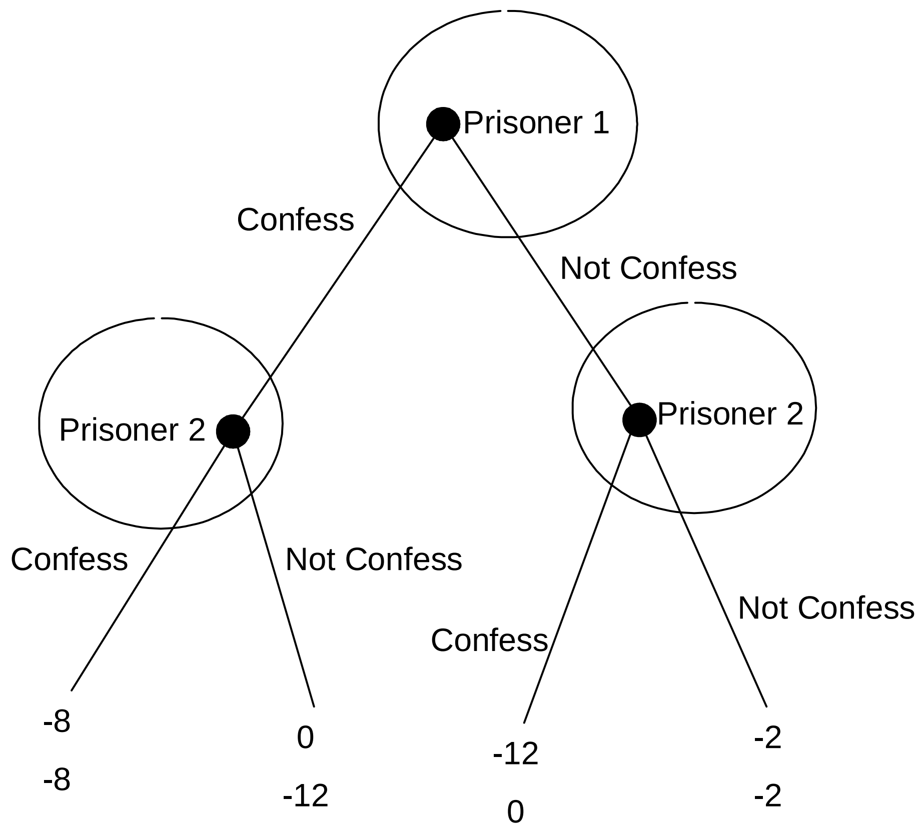


Figure 2: Sequential move Prisoner's Dilemma game.

true, Prisoner 2 now has two information sets – he needs to specify an action at EVERY information set. So Prisoner 2's four strategies are:

1. Confess if Prisoner 1 confesses, Confess if Prisoner 1 does not confess (essentially Always Confess)
2. Confess if Prisoner 1 confesses, Not Confess if Prisoner 1 does not confess (essentially play the same as Prisoner 1)
3. Not Confess if Prisoner 1 confesses, Confess if Prisoner 1 does not confess (essentially play the opposite of Prisoner 1)
4. Not Confess if Prisoner 1 confesses, Not Confess if Prisoner 1 does not confess (essentially Always Not Confess)

Given that Prisoner 2 has four strategies, we now have a 2x4 (or a 4x2) matrix representation of the sequential Prisoner's Dilemma.

		Prisoner 1	
		Confess (C)	Not Confess (NC)
	Confess if P1 C, Confess if P1 NC	$-8, -8$	$0, -12$
Prisoner 2	Confess if P1 C, Not Confess if P1 NC	$-8, -8$	$-2, -2$
	Not Confess if P1 C, Confess if P1 NC	$-12, 0$	$0, -12$
	Not Confess if P1 C, Not Confess if P1 NC	$-12, 0$	$-2, -2$

This normal form representation illustrates what a strategy is for a player in a sequential game. Finding the NE, we see that the only pure strategy NE to the sequential game is that Prisoner 1 chooses Confess and Prisoner 2 chooses "Confess if P1 C, Confess if P1 NC" or "Always Confess". Thus, there are no NE that rely on noncredible threats in this game, so there is no difference in the OUTCOME in the sequential and the simultaneous Prisoner's Dilemma games. But the NE are NOT the same because Prisoner 2 has a different strategy set in the two games.

1.2 Finding "credible" NE in an extensive form game

Because we can represent any extensive form game in strategic form, why bother with the game tree? The game tree allows us to use a concept called backward induction to find those NE which are sequentially rational and eliminate those NE which are not. In the Entrant, Incumbent game the NE of Stay Out, Fight is NOT sequentially rational while the NE of Enter, Accommodate is. Sequentially rational means that all players are choosing optimally at any point in the tree. Backward induction says that to find theses sequentially rational NE one starts at the end of the tree (the terminal nodes) and works backwards, choosing optimally at each decision node and "eliminating" the branches of the tree that are not chosen. The NE found by this method is known as the subgame perfect Nash Equilibrium (SPNE – do not confuse with pure strategy Nash Equilibrium, PSNE). What is a subgame?

Definition 1 *A subgame of an extensive form game Γ_E is a subset of the game having the following properties*

1. *it begins with an information set containing a single decision node, contains all the decision nodes that are successor nodes of this node, and contains only those nodes*
2. *If decision node x is in the subgame, then every $x' \in H(x)$ is also, where $H(x)$ is the information set that contains x (there are no broken information sets)*

Now that we know what a subgame is, we can define a SPNE. Note that the entire game tree is a subgame.

Definition 2 *A profile of strategies $\sigma = (\sigma_1, ..., \sigma_I)$ in an I-player extensive form game Γ_E is a subgame perfect Nash Equilibrium if it induces a Nash Equilibrium in every subgame of Γ_E*

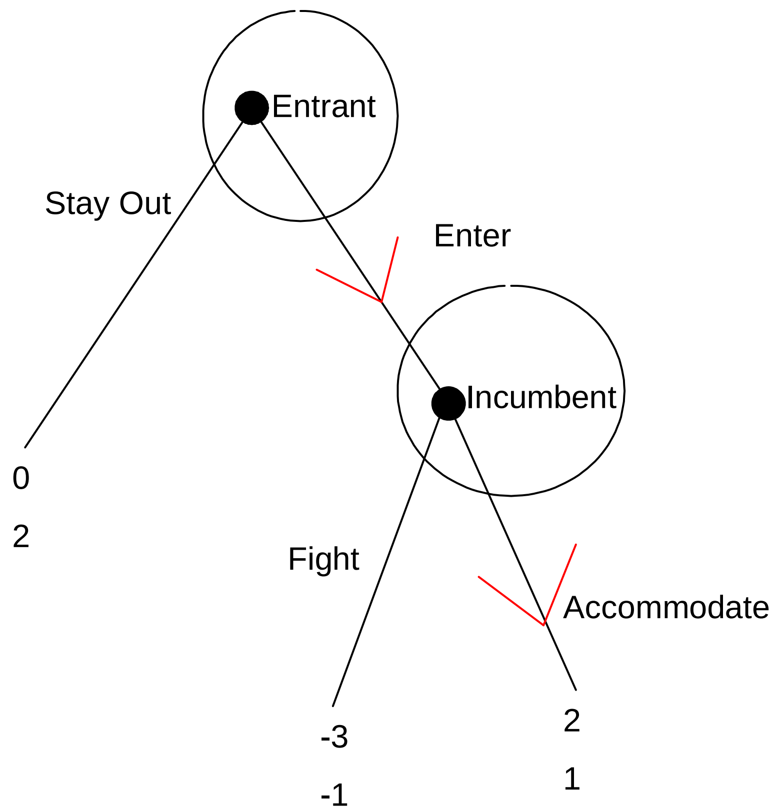


Figure 3: Illustration of SPNE of the Entrant, Incumbent game.

Basically, if we look at each of the subgames in a game tree we want players to be playing a NE in each of those subgames. This is why backward induction works as a solution technique. We start at the smallest subgames and work towards the largest subgame (the entire game) finding optimal choices at each of the smaller subgames. Consider the Entrant, Incumbent game again. We already know that there are 3 NE (2 PSNE and 1 MSNE). However, there is only 1 subgame perfect NE (SPNE), which is the Entrant Enters and the Incumbent Accommodates. To see this, start at the smallest subgame, which is the Incumbent's decision node and the two branches extending from it. A rational Incumbent would choose to Accommodate in this subgame because $1 > -1$. Knowing this, the Entrant can now eliminate the "Fight" branch from the tree, because a rational Incumbent will not play this strategy if the game gets to this point. The Entrant now has to choose between Stay Out, which has a payoff of 0 for the Entrant, and Enter, which results in a payoff of 2 for the Entrant. Because $2 > 0$, the Entrant will choose Enter.The difference between Figure 3 and Figure 1.1 is that the subgame perfect choices made by the players now have red arrows indicating the choices the players would make. This is standard notation for indicating choices in an extensive form game (the arrows at least, not necessarily the red color). We can also find the SPNE of the sequential prisoner's dilemma game using the same techniques, and it will show that Prisoner 1 Confesses, Prisoner 2 Always Confesses is the SPNE of that game. Note that all SPNE are NE, but not that all NE are SPNE. As for existence of NE and SPNE in sequential games, we have two propositions.

Proposition 3 *(Zermelo's Theorem) Every finite game of perfect information Γ_E has a pure strategy Nash Equilibrium that can be derived through backward induction. Moreover, if no player has the same payoffs at any two terminal nodes, then there is a unique Nash Equilibrium that can be derived in this manner.*

Proposition 4 *Every finite game of perfect information Γ_E has a pure strategy subgame perfect Nash Equilibrium that can be derived through backward induction. Moreover, if no player has the same payoffs at* *any two terminal nodes, then there is a unique subgame perfect Nash Equilibrium that can be derived in this manner.*

Note that we need perfect information in order to have these propositions hold, not just common knowledge. Perfect information requires that all information sets contain a single decision node. Thus, the simultaneous move Matching Pennies game and the simultaneous move Prisoner's Dilemma game do not have perfect information, so there is no guarantee that there is a PSNE to either of those games (there might be though).

2 Repeated Interactions

We have studied simultaneous and sequential games, and those games have essentially been one-shot games in nature. One-shot games are a starting point for the discussion of games, but many "games" are played repeatedly between players. Consider the following game:

		Player 2	
		Defect	Cooperate
Player 1	Defect	8, 8	32, 4
	Cooperate	4, 32	25, 25

Note that Defect is a dominant strategy for both players, but that if they could agree to Cooperate they would earn more. Note that this game is simply a Prisoner's Dilemma without the negative payoffs and the underlying story. What if the players played this game multiple, but finite, times? What would the SPNE of the repeated game be? Consider two repetitions of the game, and recall that a SPNE must induce NE in every subgame. Start with the ending subgames. Since these are simply Prisoner's Dilemmas the players must both choose Defect in order to induce NE of the subgame. Now the players have a choice of receiving either 16 and 40 (if they choose Defect) or 12 and 25 (if they choose Cooperate). Since $16 > 12$ and $40 > 25$ the players will both choose Defect in the initial play of the game. Thus, the SPNE of this twice repeated Prisoner's Dilemma is for Player 1 to always choose Defect at any decision node and for Player 2 to always choose Defect at any decision node. Regardless of how many times this game is *finitely* repeated, the only SPNE of this game will be for both of the players to choose Defect at any decision node. Thus, any attempt at cooperation in a finitely repeated Prisoner's Dilemma should unravel according to SPNE.

One might think that the players should be able to cooperate if they are playing this game repeatedly. After all, if the game is played 1000 times it is a lot better to receive 25 each period than it is to receive 8 each period. But SPNE is what it is for this game. However, what if the game was repeated *infinitely*? The first question to ask is if infinite repetition even makes sense given that the lives of humans are finite (at least to the best of our knowledge). Consider an "economic agent" that is a corporation. The corporation may be infinitely lived as it passes from one owner to the next. Second, while human lives may be finite there is (usually) some uncertainty as to when one's life will end. We can show that having an uncertain endpoint is consistent with infinite repetition of a game. Finally, and this answers the question of why infinite repetition before discussing the concept, people DO cooperate with one another on a daily basis. Infinite repetition of a one-shot game like the Prisoner's Dilemma will allow the (Cooperate, Cooperate) outcome to occur at every repetition of the game as part of a viable SPNE of the game. Be warned, however, that infinite repetition of the one-shot Prisoner's Dilemma also allows the (Defect, Defect) outcome to occur at every repetition of the game as part of a viable SPNE of the game. Thus, while infinite repetition will allow cooperation as part of the SPNE, infinite repetition also allows for a multiplicity of equilibria. This result is what is known as the "embarrassment of riches" of infinitely repeated games. Recall from earlier discussions that economists like to answer two questions when discussing the concept of equilibrium: Does an equilibrium exist and is it unique? We focus on showing the sufficient conditions for equilibrium to exist and skip the uniqueness question when discussing infinitely repeated games. There is a third question which some of you may be interested in: Among the multiple equilibria that exist, how do the players choose one of them? This question is essentially the basic one of evolutionary economics, which attempts to move economics from a physics framework to a biological one. While this approach seems novel, it has roots dating back at least to Alfred Marshall, who wrote the primary economics text (Principles of Economics – not a very clever title, but to the point) around the turn of the century – the 20^{th} century (published in 1890).

2.1 Evaluating strategies in infinite games

In order to evaluate strategies in infinite games it will be necessary to add a particular parameter to the discussion. The parameter added will be the player's discount factor, δ. It is assumed that $\delta \in [0, 1)$, and that players have exponential discounting. All that exponential discounting means is that a payoff one time period from today is discounted at δ and a payoff two time periods from today is discounted at δ^2, etc. Thus, a player's payoff stream from the infinite game would look like:

$$\delta^0\Pi_0 + \delta^1\Pi_1 + \delta^2\Pi_2 + \delta^3\Pi_3 + ...$$

where Π_k denotes the player's payoff in each period k. The $\delta \in [0, 1)$ assumption will be justified shortly.[3] It is typically assumed that players (and people in general) prefer \$1 today to \$1 tomorrow, and \$1 tomorrow to \$1 two days from now. Thus, the sooner a player receives a payoff the less the payoff is discounted. Why add this discount factor? Well, if we do not have a discount factor then the players' payoffs from following ANY strategy (assuming that there are no negative payoffs that the player could incur) of an infinite game would be infinite, which is not very interesting. This possibility of an infinite payoff is also why we assume that $\delta < 1$ rather than $\delta \leq 1$. If $\delta = 1$, then a player weights all payoffs equally regardless of the time period, and this leads to an infinite payoff. If $\delta = 0$, then the player will only care about the current period. As δ moves closer to 1, the player places more weight on future periods. It is possible to motivate this discount factor from a present value context, which should make $\delta = \frac{1}{1+r}$, where r is "the interest rate." Thus, if $r = 0.05$, then $\delta \approx 0.95$. All this says is that getting \$1 one period from today is like getting 95 cents today, and getting \$1 two periods from today is like getting 90.7 cents today. While this interpretation of the discount factor is the most closely linked to economic behavior, we will not assume that the discount factor is directly related to the interest rate, but that it is simply a parameter that states how players value payoffs over time.

Now, suppose that players 1 and 2 use the following strategies:

Player 1 chooses Cooperate in the initial period (at time $t = 0$) and continues to choose Cooperate at every decision node unless he observes that Player 2 has chosen Defect. If Player 1 ever observes Player 2 choosing Defect then Player 1 will choose Defect at every decision node after that defection. Player 2's strategy is the same. These strategies call for Cooperation at every decision node until a Defection is observed and then Defection at every decision node after Defection is observed. Note that this is a *potential* SPNE because it is a set of strategies that specifies an action at every decision node of the game. The question then becomes whether or not this set of strategies is an SPNE of the game. Recall that a strategy profile is an SPNE if and only if it specifies a NE at every subgame. Although each subgame of this game has a distinct history of play, all subgames have an identical structure. Each subgame is an infinite Prisoner's Dilemma exactly like the game as a whole. To show that these strategies are SPNE, we must show that after any previous history of play the strategies specified for the remainder of the game are NE.

Consider the following two possibilities:

1. A subgame that contains a deviation from the Cooperate, Cooperate outcome somewhere prior to the play of the subgame

2. A subgame that does not contain a deviation from the Cooperate, Cooperate outcome

If a subgame contains a deviation then the players will both choose Defect, Defect for the remainder of the game. Because this set of actions is the NE to the one-shot version (or stage game) of the Prisoner's Dilemma, it induces a NE at every subgame. Thus, the "Defect if defection has been observed" portion of the suggested strategy induces NE at every subgame.

Now, for the more difficult part. Suppose that the players are at a subgame where no previous defection has occurred. Consider the potential of deviation from the proposed strategy in period $\tau \geq t$, where t is the current period. If Player 2 chooses Defect in period τ he will earn $\delta^\tau \Pi^{Deviate} + \delta^\tau \sum_{i=1}^{\infty} \delta^i \Pi^D$ for the remainder of the game, where $\Pi^{Deviate}$ is Player 2's payoff from deviating and Π^D is his payoff each period from the (Defect, Defect) outcome. If Player 2 chooses to follow the proposed strategy, then he

[3] The exponential discounting assumption is used because it allows for time consistent preferences. Hyperbolic discounting is another type of discounting that has been suggested as consistent with choices made by individuals in experiments, although hyperbolic discounting does not necessarily lead to time consistent preferences.

will earn $\delta^\tau \sum_{i=0}^{\infty} \delta^i \Pi^C$, where Π^C is his payoff from the (Cooperate, Cooperate) outcome. The question then becomes under what conditions will the payoff from deviating be greater than that from the payoff of following the proposed strategy. To find the condition simply set up the inequality:

$$\delta^\tau \Pi^{Deviate} + \delta^\tau \sum_{i=1}^{\infty} \delta^i \Pi^D \geq \delta^\tau \sum_{i=0}^{\infty} \delta^i \Pi^C$$

We can cancel out the δ^τ terms to obtain:[4]

$$\Pi^{Deviate} + \sum_{i=1}^{\infty} \delta^i \Pi^D \geq \sum_{i=0}^{\infty} \delta^i \Pi^C$$

Now, using results on series from calculus, we have:

$$\Pi^{Deviate} + \frac{\delta}{1-\delta}\Pi^D \geq \frac{1}{1-\delta}\Pi^C$$

Now, we can substitute in for $\Pi^{Deviate}, \Pi^D$, and Π^C from our game to find:

$$32 + 8\frac{\delta}{1-\delta} \geq 25\frac{1}{1-\delta}$$

Or:

$$\begin{aligned} 32 - 32\delta + 8\delta &\geq 25 \\ 7 - 24\delta &\geq 0 \\ 7 &\geq 24\delta \\ \frac{7}{24} &\geq \delta \end{aligned}$$

Thus, choosing to deviate from the proposed strategy only provides a higher payoff if $\delta \leq \frac{7}{24}$, so that continuing to cooperate is a best response if $\delta \geq \frac{7}{24}$. The discount factor will be a key factor in determining whether or not a proposed equilibrium is an SPNE. In fact, when looking at infinitely repeated games, it is best to have a particular strategy in mind and then check to see what the necessary conditions are for it to be a SPNE, given the multiplicity of equilibria.

Are there other SPNE to the game? Consider a modified version of the game:

		Player 2	
		Defect	Cooperate
Player 1	Defect	8, 8	80, 4
	Cooperate	4, 80	25, 25

The only change in this game is that the payoff of 32 that the player received from Defecting when the other player Cooperates has been changed to 80. We can show that both players using a strategy of cooperating until a defection occurs (the same proposed strategy from before) is a SPNE if:

$$80 + 8\frac{\delta}{1-\delta} \geq 25\frac{1}{1-\delta}$$

or $\delta \geq \frac{55}{72}$. Thus, if both players are sufficiently patient then the proposed strategy is still an SPNE. Note that the discount factor increased in this example because the payoff to deviating increased. But, is there a strategy that yields higher payoffs? What if the following strategies were used by players 1 and 2:

If no deviation has occurred, Player 1 chooses Defect in all even time periods and chooses Cooperate in all odd time periods. If a deviation occurs Player 1 always chooses Defect.

If no deviation has occurred, Player 2 chooses Cooperate in all even time periods and chooses Defect in all odd time periods. If a deviation occurs Player 2 always chooses Defect.

A deviation (from player 1's perspective) occurs when Player 2 chooses Defect in an even time period. A deviation (from player 2's perspective) occurs when Player 1 chooses Defect in an odd time period. Note that we start the game at time $t = 0$, so that Player 1 receives 80 first.

[4] This canceling out of the δ^τ terms typically leads to the assumption that if deviation is going to occur in an infinitely repeated game it will occur in the first time period. I proceed under this assumption in later examples.

Look at what this strategy would do. It would cause the outcome of the game to alternate between the $(Defect, Cooperate)$ and $(Cooperate, Defect)$ outcomes, giving the players alternating periods of payoffs of 80 and 4, as opposed to 25 each period using the “cooperate until defect is observed, then always defect” strategy. On average (and ignoring discounting for a moment), each player would receive 42 per period under this new strategy and only 25 per period under the old. Is the new strategy a SPNE? We should check for both players now that they are receiving different amounts of payoffs in different periods.

For Player 1:

$$\begin{aligned} \Pi^{Deviate} &= 80 + \textstyle\sum_{i=1}^{\infty} \delta^i 8 \\ \Pi^C &= \textstyle\sum_{i=0}^{\infty} \delta^{2i} 80 + \sum_{i=0}^{\infty} \delta^{2i+1} 4 \end{aligned}$$

If $\Pi^C \geq \Pi^{Deviate}$ then Player 1 will choose NOT to deviate:

$$\begin{aligned} 80\frac{1}{1-\delta^2} + 4\frac{\delta}{1-\delta^2} &\geq 80 + 8\frac{\delta}{1-\delta} \\ 80 + 4\delta &\geq 80\left(1-\delta^2\right) + 8\delta\left(1+\delta\right) \\ 4\delta &\geq -80\delta^2 + 8\delta + 8\delta^2 \\ 72\delta^2 - 4\delta &\geq 0 \\ 18\delta - 1 &\geq 0 \\ \delta &\geq \frac{1}{18} \end{aligned}$$

This is true, for any $\delta \geq \frac{1}{18}$.

For Player 2:

$$\begin{aligned} \Pi^{Deviate} &= \textstyle\sum_{i=0}^{\infty} \delta^i 8 \\ \Pi^C &= \textstyle\sum_{i=0}^{\infty} \delta^{2i} 4 + \sum_{i=0}^{\infty} \delta^{2i+1} 80 \end{aligned}$$

If $\Pi^C \geq \Pi^{Deviate}$ then Player 2 will choose NOT to deviate:

$$\begin{aligned} 4\frac{1}{1-\delta^2} + 80\frac{\delta}{1-\delta^2} &\geq 8\frac{1}{1-\delta} \\ 4 + 80\delta &\geq 8 + 8\delta \\ 72\delta &\geq 4 \\ \delta &\geq \frac{1}{18} \end{aligned}$$

Thus, both players need to have a discount factor greater than or equal to $\frac{1}{18}$ to support this strategy. Note that this discount factor is much lower than the one needed to support the “cooperate until defect is observed, then always defect” strategy. However, it also illustrates the “embarrassment of riches” of infinitely repeated games.

2.2 Some formalities

We will now formalize some of these concepts. The focus is on 2-player games. In the one-period stage game, each player i has a compact strategy set S_i, where $q_i \in S_i$ is a particular feasible action for player i.

Let $q = (q_1, q_2)$ and $S = S_1 x S_2$.

Let $\pi_i(q_i, q_j)$ be player i's payoff function.

Let $\widehat{\pi}_i(q_j) = Max_{q \in S_i} \pi_i(q_i, q_j)$ be player i's one period best response payoff given that his rival chooses q_j.

Let $q^* = (q_1^*, q_1^*)$ denote the unique PSNE to the one-period stage game (a simplifying assumption).

A pure strategy in this game for player i, s_i, is a sequence of functions, $\{s_{it}(\cdot)\}_{t=1}^{\infty}$ mapping from the history of previous action choices (denoted H_{t-1}) to a player's action choice in period t, $s_{it}(H_{t-1}) \in S_i$.

The set of all pure strategies for player i is denoted Σ_i, and $s = (s_1, s_2) \in \Sigma_1 x \Sigma_2$ is a profile of pure strategies for the players.

Any pure strategy profile $s = (s_1, s_2)$ induces an outcome path $Q(s)$, which is an infinite sequence of actions $\{q_t = (q_{1t}, q_{2t})\}_{t=1}^{\infty}$ that will actually be played when the players follow strategies s_1 and s_2.

Player i's discounted payoff from outcome path Q is denoted by $v_i(Q) = \sum_{t=0}^{\infty} \delta^t \pi_i(q_{1+t})$.

Player i's average payoff from outcome path Q is $(1-\delta) v_i(Q)$.

Player i's discounted continuation payoff from some point t onward is $v_i(Q, t) = \sum_{\tau=0}^{\infty} \delta^{\tau} \pi_i(q_{t+\tau})$.

We already know that the strategies that call for player i to play the stage game NE q_i^* in every period, regardless of the prior history, constitute an SPNE for any $\delta < 1$.

2.2.1 Nash reversion and the Nash reversion Folk Theorem

Nash reversion is essentially the "punishment" we have been discussing – if one player fails to "cooperate", the other player "punishes" by reverting to the stage game NE. It was well-known that this was a solution to the infinitely repeated game before someone decided to write it down, hence the term "Folk Theorem".

Definition 5 *A strategy profile $s = (s_1, s_2)$ in an infinitely repeated game is one of Nash reversion if each player's strategy calls for playing some outcome path Q until someone deviates and playing the stage game NE $q^* = (q_1^*, q_2^*)$ thereafter.*

Lemma 6 *A Nash reversion strategy profile that calls for playing outcome path $Q = \{q_{1t}, q_{2t}\}_{t=1}^{\infty}$ prior to any deviation is a SPNE if and only if*

$$\widehat{\pi}_i(q_{jt}) + \frac{\delta}{1-\delta}\pi_i(q_i^*, q_j^*) \leq v_i(Q, t)$$

where $j \neq i$ for all t and $i = 1, 2$.

This formalizes what we have already been discussing in the context of the Prisoner's Dilemma game.

Proposition 7 *Consider an infinitely repeated game with $\delta > 0$ and $S_i \subset \mathbb{R}$ for $i = 1, 2$. Suppose also that $\pi_i(q)$ is differentiable at $q^* = (q_1^*, q_2^*)$ with $\partial \pi_i(q_1^*, q_2^*)/\partial q_j \neq 0$ for $j \neq i$ and $i = 1, 2$. Then there is some $q' = (q_1', q_2')$ with $[\pi_1(q_1', q_2'), \pi_2(q_1', q_2')] >> [\pi_1(q_1^*, q_2^*), \pi_2(q_1^*, q_2^*)]$ whose infinite repetition is the outcome path of an SPNE that uses Nash reversion.*

This proposition states that with continuous strategy sets and differentiable payoff functions, as long as there is some possibility for joint improvement in payoffs around the stage game NE some cooperation can be sustained.

Proposition 8 *Suppose that outcome path Q can be sustained as an SPNE outcome path using Nash reversion when the discount factor is δ. Then it can be so sustained for any $\delta' \geq \delta$.*

Hopefully that proposition is self-explanatory ...

Proposition 9 *For any pair of actions $q = (q_1, q_2)$ such that $\pi_i(q_1, q_2) > \pi_i(q_1^*, q_2^*)$ for $i = 1, 2$ there exists a $\underline{\delta} < 1$ such that, for all $\delta > \underline{\delta}$, infinite repetition of $q = (q_1, q_2)$ is the outcome path of an SPNE using Nash reversion strategies.*

This proposition is the most important one as it states that any stationary outcome path that gives each player a discounted payoff that exceeds the payoff arising from infinite repetition of the stage game NE can be sustained as an SPNE if δ is sufficiently close to 1. Note that this proposition only applies to stationary paths. It is possible to extend the argument to include non-stationary paths with average payoff vectors greater than the stage game NE.

The text also discusses some other propositions which show that payoff vectors LESS than the stage game NE payoff vector can be supported using punishment strategies that are harsher than Nash reversion. Essentially, Player 1 threatens to punish Player 2 by forcing Player 2 to accept the minimum amount he possibly could be forced to accept. In the simple Prisoner's Dilemma we have been discussing this is the same as Nash reversion because Player 2 could not be forced to accept anything less than the stage game NE.